

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: TANABE, Tadashi
- (ii) TITLE OF INVENTION: PROSTACYCLIN SYNTHASE DERIVED FROM HUMAN
- (iii) NUMBER OF SEQUENCES: 17
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
 - (B) STREET: 2100 Pennsylvania Avenue, N.W.
 - (C) CITY: Washington
 - (D) STATE: D.C.
 - (E) COUNTRY: USA
 - (F) ZIP: 20037
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/578,709
 - (B) FILING DATE: 28-DEC-1995
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: PCT/JP95/00838
 - (B) FILING DATE: 27-APR-1995
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: JP 114316/1994
 - (B) FILING DATE: 28-APR-1994
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Gubinsky, Louis
 - (B) REGISTRATION NUMBER: 24,835
 - (C) REFERENCE/DOCKET NUMBER: Q40439
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (202)293-7060
 - (B) TELEFAX: (202)293-7860

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 16 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PRIMER/SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GGGACAAGGA CCACAT

16

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PRIMER/SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CAAAAGTCGC CTGTGGAAGC

20

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PRIMER/SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CACAGGCGAC TTTTGACA

18

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PRIMER/SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

TGCCTGCATC TCCTCTGA

18

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PRIMER/SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GACTCGAGTC GACATCGATT TTTTTTTTTT TTTT

35

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PRIMER/SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GACTCGAGTC GACATCG

17

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 48 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PRIMER/SYNTHETIC RNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CUACUACUAC UAGGCCACGC GTCGACTAGT ACGGGNNGG NNGGGNNG

48

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 183 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens
- (B) STRAIN: lambda hPGIS141

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..183

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GGG GAC AAG GAC CAC ATG TGC AGT GTC AAA AGT CGC CTG TGG AAG CTG	48
Gly Asp Lys Asp His Met Cys Ser Val Lys Ser Arg Leu Trp Lys Leu	
1 5 10 15	
CTA TCC CCA GCC AGG CTG GCC AGG CGG GCC CAC CGG AGC AAA TGG CTG	96
Leu Ser Pro Ala Arg Leu Ala Arg Arg Ala His Arg Ser Lys Trp Leu	
20 25 30	
GAG AGT TAC CTG CTG CAC CTG GAG GAG ATG GGT GTG TCA GAG GAG ATG	144
Glu Ser Tyr Leu Leu His Leu Glu Glu Met Gly Val Ser Glu Glu Met	
35 40 45	
CAG GCA CGG GCC CTG GTG CTG CAG CTG TGG GCC ACA CAG	183
Gln Ala Arg Ala Leu Val Leu Gln Leu Trp Ala Thr Gln	
50 55 60	

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 61 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Gly Asp Lys Asp His Met Cys Ser Val Lys Ser Arg Leu Trp Lys Leu	
1 5 10 15	
Leu Ser Pro Ala Arg Leu Ala Arg Arg Ala His Arg Ser Lys Trp Leu	
20 25 30	

Glu Ser Tyr Leu Leu His Leu Glu Glu Met Gly Val Ser Glu Glu Met
 35 40 45

Gln Ala Arg Ala Leu Val Leu Gln Leu Trp Ala Thr Gln
 50 55 60

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 792 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens
- (B) STRAIN: pHPGIS36

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 80..790

(ix) FEATURE:

- (A) NAME/KEY: mat_peptide
- (B) LOCATION: 80..790

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CTACTACTAC TAGGCCACGC GTCGACTAGT ACGGGGGGGG GGGGGGGGGG GCAGCCCCGC	60
CAGCCCCGCC AGCCCCGCG ATG GCT TGG GCC GCG CTC CTC GGC CTC CTG GCC	112
Met Ala Trp Ala Ala Leu Leu Gly Leu Leu Ala	
1 5 10	
GCA CTG TTG CTG CTG CTG CTA CTG AGC CGC CGC CGC ACG CGG CGA CCT	160
Ala Leu Leu Leu Leu Leu Leu Leu Ser Arg Arg Arg Thr Arg Arg Pro	
15 20 25	
GGT GAG CCT CCC CTG GAC CTG GGC AGC ATC CCC TGG TTG GGG TAT GCC	208
Gly Glu Pro Pro Leu Asp Leu Gly Ser Ile Pro Trp Leu Gly Tyr Ala	
30 35 40	
TTG GAC TTT GGA AAA GAT GCT GCC AGC TTC CTC ACG AGG ATG AAG GAG	256
Leu Asp Phe Gly Lys Asp Ala Ala Ser Phe Leu Thr Arg Met Lys Glu	
45 50 55	
AAG CAC GGT GAC ATC TTT ACT ATA CTG GTT GGG GGC AGG TAT GTC ACC	304
Lys His Gly Asp Ile Phe Thr Ile Leu Val Gly Gly Arg Tyr Val Thr	
60 65 70 75	
GTT CTC CTG GAC CCA CAC TCC TAC GAC GCG GTG GTG TGG GAG CCT CGC	352
Val Leu Leu Asp Pro His Ser Tyr Asp Ala Val Val Trp Glu Pro Arg	
80 85 90	

ACC AGG CTC GAC TTC CAT GCC TAT GCC ATC TTC CTC ATG GAG AGG ATT	400
Thr Arg Leu Asp Phe His Ala Tyr Ala Ile Phe Leu Met Glu Arg Ile	
95 100 105	
TTT GAT GTG CAG CTT CCA CAT TAC AGC CCC AGT GAT GAA AAG GCC AGG	448
Phe Asp Val Gln Leu Pro His Tyr Ser Pro Ser Asp Glu Lys Ala Arg	
110 115 120	
ATG AAA CTG ACT CTT CTC CAC AGA GAG CTC CAG GCA CTC ACA GAA GCC	496
Met Lys Leu Thr Leu Leu His Arg Glu Leu Gln Ala Leu Thr Glu Ala	
125 130 135	
ATG TAT ACC AAC CTC CAT GCA GTG CTG TTG GGC GAT GCT ACA GAA GCA	544
Met Tyr Thr Asn Leu His Ala Val Leu Leu Gly Asp Ala Thr Glu Ala	
140 145 150 155	
GGC AGT GGC TGG CAC GAG ATG GGT CTC CTC GAC TTC TCC TAC AGC TTC	592
Gly Ser Gly Trp His Glu Met Gly Leu Leu Asp Phe Ser Tyr Ser Phe	
160 165 170	
CTG CTC AGA GCC GGC TAC CTG ACT CTT TAC GGA ATT GAG GCG CTG CCA	640
Leu Leu Arg Ala Gly Tyr Leu Thr Leu Tyr Gly Ile Glu Ala Leu Pro	
175 180 185	
CGC ACC CAT GAA AGC CAG GCC CAG GAC CGC GTC CAC TCA GCT GAT GTC	688
Arg Thr His Glu Ser Gln Ala Gln Asp Arg Val His Ser Ala Asp Val	
190 195 200	
TTC CAC ACC TTT CGC CAG CTC GAC CGG CTG CTC CCC AAA CTG GCC CGT	736
Phe His Thr Phe Arg Gln Leu Asp Arg Leu Leu Pro Lys Leu Ala Arg	
205 210 215	
GGC TCC CTG TCA GTG GGG GAC AAG GAC CAC ATG TGC AGT GTC AAA AGT	784
Gly Ser Leu Ser Val Gly Asp Lys Asp His Met Cys Ser Val Lys Ser	
220 225 230 235	
CGC CTG TG	792
Arg Leu	

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 237 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Met	Ala	Trp	Ala	Ala	Leu	Leu	Gly	Leu	Leu	Ala	Ala	Leu	Leu	Leu	Leu
1					5				10					15	
Leu	Leu	Leu	Ser	Arg	Arg	Arg	Thr	Arg	Arg	Pro	Gly	Glu	Pro	Pro	Leu

20	25	30
Asp Leu Gly Ser Ile Pro Trp Leu Gly Tyr Ala Leu Asp Phe Gly Lys		
35	40	45
Asp Ala Ala Ser Phe Leu Thr Arg Met Lys Glu Lys His Gly Asp Ile		
50	55	60
Phe Thr Ile Leu Val Gly Gly Arg Tyr Val Thr Val Leu Leu Asp Pro		
65	70	75
His Ser Tyr Asp Ala Val Val Trp Glu Pro Arg Thr Arg Leu Asp Phe		
85	90	95
His Ala Tyr Ala Ile Phe Leu Met Glu Arg Ile Phe Asp Val Gln Leu		
100	105	110
Pro His Tyr Ser Pro Ser Asp Glu Lys Ala Arg Met Lys Leu Thr Leu		
115	120	125
Leu His Arg Glu Leu Gln Ala Leu Thr Glu Ala Met Tyr Thr Asn Leu		
130	135	140
His Ala Val Leu Leu Gly Asp Ala Thr Glu Ala Gly Ser Gly Trp His		
145	150	155
Glu Met Gly Leu Leu Asp Phe Ser Tyr Ser Phe Leu Leu Arg Ala Gly		
165	170	175
Tyr Leu Thr Leu Tyr Gly Ile Glu Ala Leu Pro Arg Thr His Glu Ser		
180	185	190
Gln Ala Gln Asp Arg Val His Ser Ala Asp Val Phe His Thr Phe Arg		
195	200	205
Gln Leu Asp Arg Leu Leu Pro Lys Leu Ala Arg Gly Ser Leu Ser Val		
210	215	220
Gly Asp Lys Asp His Met Cys Ser Val Lys Ser Arg Leu		
225	230	235

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1296 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens
- (B) STRAIN: pHPGIS135

(ix) FEATURE:

- (A) NAME/KEY: CDS

(B) LOCATION: 3..827

(ix) FEATURE:

(A) NAME/KEY: mat_peptide

(B) LOCATION: 3..827

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GG GAC AAG GAC CAC ATG TGC AGT GTC AAA AGT CGC CTG TGG AAG CTG	47
Asp Lys Asp His Met Cys Ser Val Lys Ser Arg Leu Trp Lys Leu	
1 5 10 15	
CTA TCC CCA GCC AGG CTG GCC AGG CGG GCC CAC CGG AGC AAA TGG CTG	95
Leu Ser Pro Ala Arg Leu Ala Arg Arg Ala His Arg Ser Lys Trp Leu	
20 25 30	
GAG AGT TAC CTG CTG CAC CTG GAG GAG ATG GGT GTG TCA GAG GAG ATG	143
Glu Ser Tyr Leu Leu His Leu Glu Glu Met Gly Val Ser Glu Glu Met	
35 40 45	
CAG GCA CGG GCC CTG GTG CTG CAG CTG TGG GCC ACA CAG GGG AAT ATG	191
Gln Ala Arg Ala Leu Val Leu Gln Leu Trp Ala Thr Gln Gly Asn Met	
50 55 60	
GGT CCC GCT GCC TTC TGG CTC CTG CTC TTC CTT CTC AAG AAT CCT GAA	239
Gly Pro Ala Ala Phe Trp Leu Leu Leu Phe Leu Leu Lys Asn Pro Glu	
65 70 75	
GCC CTG GCT GCT GTC CGC GGA GAG CTC GAG AGT ATC CTT TGG CAA GCG	287
Ala Leu Ala Ala Val Arg Gly Glu Leu Glu Ser Ile Leu Trp Gln Ala	
80 85 90 95	
GAG CAG CCT GTC TCG CAG ACG ACC ACT CTC CCA CAG AAG GTT CTA GAC	335
Glu Gln Pro Val Ser Gln Thr Thr Thr Leu Pro Gln Lys Val Leu Asp	
100 105 110	
AGC ACA CCT GTG CTT GAT AGC GTG CTG AGT GAG AGC CTC AGG CTT ACA	383
Ser Thr Pro Val Leu Asp Ser Val Leu Ser Glu Ser Leu Arg Leu Thr	
115 120 125	
GCT GCC CCC TTC ATC ACC CGC GAG GTT GTG GTG GAC CTG GCC ATG CCC	431
Ala Ala Pro Phe Ile Thr Arg Glu Val Val Val Asp Leu Ala Met Pro	
130 135 140	
ATG GCA GAC GGG AGA GAA TTC AAC CTG CGA CGT GGT GAC CGC CTC CTC	479
Met Ala Asp Gly Arg Glu Phe Asn Leu Arg Arg Gly Asp Arg Leu Leu	
145 150 155	
CTC TTC CCC TTC CTG AGC CCC CAG AGA GAC CCA GAA ATC TAC ACA GAC	527
Leu Phe Pro Phe Leu Ser Pro Gln Arg Asp Pro Glu Ile Tyr Thr Asp	
160 165 170 175	
CCA GAG GTA TTT AAA TAC AAC CGA TTC CTG AAC CCT GAC GGA TCA GAG	575
Pro Glu Val Phe Lys Tyr Asn Arg Phe Leu Asn Pro Asp Gly Ser Glu	
180 185 190	

AAG AAA GAC TTT TAC AAG GAT GGG AAA CGG CTG AAG AAT TAC AAC ATG Lys Lys Asp Phe Tyr Lys Asp Gly Lys Arg Leu Lys Asn Tyr Asn Met 195 200 205	623
CCC TGG GGG GCG GGG CAC AAT CAC TGC CTG GGG AGG AGT TAT GCG GTC Pro Trp Gly Ala Gly His Asn His Cys Leu Gly Arg Ser Tyr Ala Val 210 215 220	671
AAC AGC ATC AAA CAA TTT GTG TTC CTT GTG CTG GTG CAC TTG GAC TTG Asn Ser Ile Lys Gln Phe Val Phe Leu Val Leu Val His Leu Asp Leu 225 230 235	719
GAG CTG ATC AAC GCA GAT GTG GAG ATC CCT GAG TTT GAC CTC AGC AGG Glu Leu Ile Asn Ala Asp Val Glu Ile Pro Glu Phe Asp Leu Ser Arg 240 245 250 255	767
TAC GGC TTC GGT CTG ATG CAG CCG GAA CAC GAC GTG CCC GTC CGC TAC Tyr Gly Phe Gly Leu Met Gln Pro Glu His Asp Val Pro Val Arg Tyr 260 265 270	815
CGC ATC CGC CCA TGACACAGGG AGCAGATGGA TCCACGTGCT CGCCTCTGCC Arg Ile Arg Pro 275	867
CAGCCTGCCC CAGCCTGCCC CAGCCTCCCA GCTTTCTGTG TGCACAGTTG GCCCGGGTGC	927
AGGTGCTAGC ATTACCACTT CCCTGCTTTT CTCCAGAAG GCTGGGTCCA GGGGAGGGAA	987
AAGCTAAGAG GGTGAACAAA GAAAAGACAT TGAAAGCTCT ATGGATTATC CACTGCAAAG	1047
TTTTCTTTCC AAAATCAGGC TTTGTCTGCT CCAATTAC CTCGTTACTC TCACCTCGTG	1107
ATATCCACAA ATGCTATTCA GATAAGGCAG AACTAGGAGT CTTCACTGCT CTGCCCCCAA	1167
CTCCCGGAGG TGTCACCTTC CTAGTTCTTA TGAGCTAGCA TGGCCCGGGC CTTATCCAGT	1227
CAAAGCGGAT GCTGGCCACA GAAAGGCCAC TCAGGATGTC CTTTGTGTCC ATCGATGTCG	1287
ACTCGAGTC	1296

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 275 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Asp Lys Asp His Met Cys Ser Val Lys Ser Arg Leu Trp Lys Leu Leu 1 5 10 15
Ser Pro Ala Arg Leu Ala Arg Arg Ala His Arg Ser Lys Trp Leu Glu 20 25 30

Ser Tyr Leu Leu His Leu Glu Glu Met Gly Val Ser Glu Glu Met Gln
 35 40 45
 Ala Arg Ala Leu Val Leu Gln Leu Trp Ala Thr Gln Gly Asn Met Gly
 50 55 60
 Pro Ala Ala Phe Trp Leu Leu Phe Leu Leu Lys Asn Pro Glu Ala
 65 70 75 80
 Leu Ala Ala Val Arg Gly Glu Leu Glu Ser Ile Leu Trp Gln Ala Glu
 85 90 95
 Gln Pro Val Ser Gln Thr Thr Thr Leu Pro Gln Lys Val Leu Asp Ser
 100 105 110
 Thr Pro Val Leu Asp Ser Val Leu Ser Glu Ser Leu Arg Leu Thr Ala
 115 120 125
 Ala Pro Phe Ile Thr Arg Glu Val Val Val Asp Leu Ala Met Pro Met
 130 135 140
 Ala Asp Gly Arg Glu Phe Asn Leu Arg Arg Gly Asp Arg Leu Leu Leu
 145 150 155 160
 Phe Pro Phe Leu Ser Pro Gln Arg Asp Pro Glu Ile Tyr Thr Asp Pro
 165 170 175
 Glu Val Phe Lys Tyr Asn Arg Phe Leu Asn Pro Asp Gly Ser Glu Lys
 180 185 190
 Lys Asp Phe Tyr Lys Asp Gly Lys Arg Leu Lys Asn Tyr Asn Met Pro
 195 200 205
 Trp Gly Ala Gly His Asn His Cys Leu Gly Arg Ser Tyr Ala Val Asn
 210 215 220
 Ser Ile Lys Gln Phe Val Phe Leu Val Leu Val His Leu Asp Leu Glu
 225 230 235 240
 Leu Ile Asn Ala Asp Val Glu Ile Pro Glu Phe Asp Leu Ser Arg Tyr
 245 250 255
 Gly Phe Gly Leu Met Gln Pro Glu His Asp Val Pro Val Arg Tyr Arg
 260 265 270
 Ile Arg Pro
 275

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1977 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Homo sapiens

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 28..1527

(ix) FEATURE:

(A) NAME/KEY: mat_peptide

(B) LOCATION: 28..1527

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

AGCCCCGCCA GCCCCGCCAG CCCC GCG ATG GCT TGG GCC GCG CTC CTC GGC	51
Met Ala Trp Ala Leu Leu Gly	
1 5	
CTC CTG GCC GCA CTG TTG CTG CTG CTG CTA CTG AGC CGC CGC CGC ACG	99
Leu Leu Ala Ala Leu Leu Leu Leu Leu Leu Ser Arg Arg Arg Thr	
10 15 20	
CGG CGA CCT GGT GAG CCT CCC CTG GAC CTG GGC AGC ATC CCC TGG TTG	147
Arg Arg Pro Gly Glu Pro Pro Leu Asp Leu Gly Ser Ile Pro Trp Leu	
25 30 35 40	
GGG TAT GCC TTG GAC TTT GGA AAA GAT GCT GCC AGC TTC CTC ACG AGG	195
Gly Tyr Ala Leu Asp Phe Gly Lys Asp Ala Ala Ser Phe Leu Thr Arg	
45 50 55	
ATG AAG GAG AAG CAC GGT GAC ATC TTT ACT ATA CTG GTT GGG GGC AGG	243
Met Lys Glu Lys His Gly Asp Ile Phe Thr Ile Leu Val Gly Gly Arg	
60 65 70	
TAT GTC ACC GTT CTC CTG GAC CCA CAC TCC TAC GAC GCG GTG GTG TGG	291
Tyr Val Thr Val Leu Leu Asp Pro His Ser Tyr Asp Ala Val Val Trp	
75 80 85	
GAG CCT CGC ACC AGG CTC GAC TTC CAT GCC TAT GCC ATC TTC CTC ATG	339
Glu Pro Arg Thr Arg Leu Asp Phe His Ala Tyr Ala Ile Phe Leu Met	
90 95 100	
GAG AGG ATT TTT GAT GTG CAG CTT CCA CAT TAC AGC CCC AGT GAT GAA	387
Glu Arg Ile Phe Asp Val Gln Leu Pro His Tyr Ser Pro Ser Asp Glu	
105 110 115 120	
AAG GCC AGG ATG AAA CTG ACT CTT CTC CAC AGA GAG CTC CAG GCA CTC	435
Lys Ala Arg Met Lys Leu Thr Leu Leu His Arg Glu Leu Gln Ala Leu	
125 130 135	
ACA GAA GCC ATG TAT ACC AAC CTC CAT GCA GTG CTG TTG GGC GAT GCT	483
Thr Glu Ala Met Tyr Thr Asn Leu His Ala Val Leu Leu Gly Asp Ala	
140 145 150	
ACA GAA GCA GGC AGT GGC TGG CAC GAG ATG GGT CTC CTC GAC TTC TCC	531

Thr	Glu	Ala	Gly	Ser	Gly	Trp	His	Glu	Met	Gly	Leu	Leu	Asp	Phe	Ser		
		155					160					165					
TAC	AGC	TTC	CTG	CTC	AGA	GCC	GGC	TAC	CTG	ACT	CTT	TAC	GGA	ATT	GAG	579	
Tyr	Ser	Phe	Leu	Leu	Arg	Ala	Gly	Tyr	Leu	Thr	Leu	Tyr	Gly	Ile	Glu		
	170					175					180						
GCG	CTG	CCA	CGC	ACC	CAT	GAA	AGC	CAG	GCC	CAG	GAC	CGC	GTC	CAC	TCA	627	
Ala	Leu	Pro	Arg	Thr	His	Glu	Ser	Gln	Ala	Gln	Asp	Arg	Val	His	Ser		
	185				190					195					200		
GCT	GAT	GTC	TTC	CAC	ACC	TTT	CGC	CAG	CTC	GAC	CGG	CTG	CTC	CCC	AAA	675	
Ala	Asp	Val	Phe	His	Thr	Phe	Arg	Gln	Leu	Asp	Arg	Leu	Leu	Pro	Lys		
				205					210					215			
CTG	GCC	CGT	GGC	TCC	CTG	TCA	GTG	GGG	GAC	AAG	GAC	CAC	ATG	TGC	AGT	723	
Leu	Ala	Arg	Gly	Ser	Leu	Ser	Val	Gly	Asp	Lys	Asp	His	Met	Cys	Ser		
		220						225					230				
GTC	AAA	AGT	CGC	CTG	TGG	AAG	CTG	CTA	TCC	CCA	GCC	AGG	CTG	GCC	AGG	771	
Val	Lys	Ser	Arg	Leu	Trp	Lys	Leu	Leu	Ser	Pro	Ala	Arg	Leu	Ala	Arg		
	235					240						245					
CGG	GCC	CAC	CGG	AGC	AAA	TGG	CTG	GAG	AGT	TAC	CTG	CTG	CAC	CTG	GAG	819	
Arg	Ala	His	Arg	Ser	Lys	Trp	Leu	Glu	Ser	Tyr	Leu	Leu	His	Leu	Glu		
	250					255					260						
GAG	ATG	GGT	GTG	TCA	GAG	GAG	ATG	CAG	GCA	CGG	GCC	CTG	GTG	CTG	CAG	867	
Glu	Met	Gly	Val	Ser	Glu	Glu	Met	Gln	Ala	Arg	Ala	Leu	Val	Leu	Gln		
	265				270					275					280		
CTG	TGG	GCC	ACA	CAG	GGG	AAT	ATG	GGT	CCC	GCT	GCC	TTC	TGG	CTC	CTG	915	
Leu	Trp	Ala	Thr	Gln	Gly	Asn	Met	Gly	Pro	Ala	Ala	Phe	Trp	Leu	Leu		
				285					290					295			
CTC	TTC	CTT	CTC	AAG	AAT	CCT	GAA	GCC	CTG	GCT	GCT	GTC	CGC	GGA	GAG	963	
Leu	Phe	Leu	Leu	Lys	Asn	Pro	Glu	Ala	Leu	Ala	Ala	Val	Arg	Gly	Glu		
			300					305					310				
CTC	GAG	AGT	ATC	CTT	TGG	CAA	GCG	GAG	CAG	CCT	GTC	TCG	CAG	ACG	ACC	1011	
Leu	Glu	Ser	Ile	Leu	Trp	Gln	Ala	Glu	Gln	Pro	Val	Ser	Gln	Thr	Thr		
	315					320						325					
ACT	CTC	CCA	CAG	AAG	GTT	CTA	GAC	AGC	ACA	CCT	GTG	CTT	GAT	AGC	GTG	1059	
Thr	Leu	Pro	Gln	Lys	Val	Leu	Asp	Ser	Thr	Pro	Val	Leu	Asp	Ser	Val		
	330					335					340						
CTG	AGT	GAG	AGC	CTC	AGG	CTT	ACA	GCT	GCC	CCC	TTC	ATC	ACC	CGC	GAG	1107	
Leu	Ser	Glu	Ser	Leu	Arg	Leu	Thr	Ala	Ala	Pro	Phe	Ile	Thr	Arg	Glu		
	345				350					355					360		
GTT	GTG	GTG	GAC	CTG	GCC	ATG	CCC	ATG	GCA	GAC	GGG	AGA	GAA	TTC	AAC	1155	
Val	Val	Val	Asp	Leu	Ala	Met	Pro	Met	Ala	Asp	Gly	Arg	Glu	Phe	Asn		
				365					370					375			
CTG	CGA	CGT	GGT	GAC	CGC	CTC	CTC	CTC	TTC	CCC	TTC	CTG	AGC	CCC	CAG	1203	
Leu	Arg	Arg	Gly	Asp	Arg	Leu	Leu	Leu	Phe	Pro	Phe	Leu	Ser	Pro	Gln		

380	385	390	
AGA GAC CCA GAA ATC TAC ACA GAC CCA GAG GTA TTT AAA TAC AAC CGA			1251
Arg Asp Pro Glu Ile Tyr Thr Asp Pro Glu Val Phe Lys Tyr Asn Arg			
395	400	405	
TTC CTG AAC CCT GAC GGA TCA GAG AAG AAA GAC TTT TAC AAG GAT GGG			1299
Phe Leu Asn Pro Asp Gly Ser Glu Lys Lys Asp Phe Tyr Lys Asp Gly			
410	415	420	
AAA CGG CTG AAG AAT TAC AAC ATG CCC TGG GGG GCG GGG CAC AAT CAC			1347
Lys Arg Leu Lys Asn Tyr Asn Met Pro Trp Gly Ala Gly His Asn His			
425	430	435	440
TGC CTG GGG AGG AGT TAT GCG GTC AAC AGC ATC AAA CAA TTT GTG TTC			1395
Cys Leu Gly Arg Ser Tyr Ala Val Asn Ser Ile Lys Gln Phe Val Phe			
445	450	455	
CTT GTG CTG GTG CAC TTG GAC TTG GAG CTG ATC AAC GCA GAT GTG GAG			1443
Leu Val Leu Val His Leu Asp Leu Glu Leu Ile Asn Ala Asp Val Glu			
460	465	470	
ATC CCT GAG TTT GAC CTC AGC AGG TAC GGC TTC GGT CTG ATG CAG CCG			1491
Ile Pro Glu Phe Asp Leu Ser Arg Tyr Gly Phe Gly Leu Met Gln Pro			
475	480	485	
GAA CAC GAC GTG CCC GTC CGC TAC CGC ATC CGC CCA TGACACAGGG			1537
Glu His Asp Val Pro Val Arg Tyr Arg Ile Arg Pro			
490	495	500	
AGCAGATGGA TCCACGTGCT CGCCTCTGCC CAGCCTGCCC CAGCCTGCCC CAGCCTCCCA			1597
GCTTTCTGTG TGCACAGTTG GCCCGGGTGC AGGTGCTAGC ATTACCACTT CCCTGCTTTT			1657
CTCCCAGAAG GCTGGGTCCA GGGGAGGGAA AAGCTAAGAG GGTGAACAAA GAAAAGACAT			1717
TGAAAGCTCT ATGGATTATC CACTGCAAAG TTTTCTTTCC AAAATCAGGC TTTGTCTGCT			1777
CCCAATTAC CTCGTTACTC TCACCTCGTG ATATCCACAA ATGCTATTCA GATAAGGCAG			1837
AACTAGGAGT CTTCACTGCT CTGCCCCCAA CTCCCGGAGG TGTCACCTTC CTAGTTCTTA			1897
TGAGCTAGCA TGGCCCGGGC CTTATCCAGT CAAAGCGGAT GCTGGCCACA GAAAGGCCAC			1957
TCAGGATGTC CTTTGTGTCC			1977

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 500 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Met	Ala	Trp	Ala	Ala	Leu	Leu	Gly	Leu	Leu	Ala	Ala	Leu	Leu	Leu	Leu	1	5	10	15
Leu	Leu	Leu	Ser	Arg	Arg	Arg	Thr	Arg	Arg	Pro	Gly	Glu	Pro	Pro	Leu	20	25	30	
Asp	Leu	Gly	Ser	Ile	Pro	Trp	Leu	Gly	Tyr	Ala	Leu	Asp	Phe	Gly	Lys	35	40	45	
Asp	Ala	Ala	Ser	Phe	Leu	Thr	Arg	Met	Lys	Glu	Lys	His	Gly	Asp	Ile	50	55	60	
Phe	Thr	Ile	Leu	Val	Gly	Gly	Arg	Tyr	Val	Thr	Val	Leu	Leu	Asp	Pro	65	70	75	80
His	Ser	Tyr	Asp	Ala	Val	Val	Trp	Glu	Pro	Arg	Thr	Arg	Leu	Asp	Phe	85	90	95	
His	Ala	Tyr	Ala	Ile	Phe	Leu	Met	Glu	Arg	Ile	Phe	Asp	Val	Gln	Leu	100	105	110	
Pro	His	Tyr	Ser	Pro	Ser	Asp	Glu	Lys	Ala	Arg	Met	Lys	Leu	Thr	Leu	115	120	125	
Leu	His	Arg	Glu	Leu	Gln	Ala	Leu	Thr	Glu	Ala	Met	Tyr	Thr	Asn	Leu	130	135	140	
His	Ala	Val	Leu	Leu	Gly	Asp	Ala	Thr	Glu	Ala	Gly	Ser	Gly	Trp	His	145	150	155	160
Glu	Met	Gly	Leu	Leu	Asp	Phe	Ser	Tyr	Ser	Phe	Leu	Leu	Arg	Ala	Gly	165	170	175	
Tyr	Leu	Thr	Leu	Tyr	Gly	Ile	Glu	Ala	Leu	Pro	Arg	Thr	His	Glu	Ser	180	185	190	
Gln	Ala	Gln	Asp	Arg	Val	His	Ser	Ala	Asp	Val	Phe	His	Thr	Phe	Arg	195	200	205	
Gln	Leu	Asp	Arg	Leu	Leu	Pro	Lys	Leu	Ala	Arg	Gly	Ser	Leu	Ser	Val	210	215	220	
Gly	Asp	Lys	Asp	His	Met	Cys	Ser	Val	Lys	Ser	Arg	Leu	Trp	Lys	Leu	225	230	235	240
Leu	Ser	Pro	Ala	Arg	Leu	Ala	Arg	Arg	Ala	His	Arg	Ser	Lys	Trp	Leu	245	250	255	
Glu	Ser	Tyr	Leu	Leu	His	Leu	Glu	Glu	Met	Gly	Val	Ser	Glu	Glu	Met	260	265	270	
Gln	Ala	Arg	Ala	Leu	Val	Leu	Gln	Leu	Trp	Ala	Thr	Gln	Gly	Asn	Met	275	280	285	
Gly	Pro	Ala	Ala	Phe	Trp	Leu	Leu	Leu	Phe	Leu	Leu	Lys	Asn	Pro	Glu	290	295	300	

Ala Leu Ala Ala Val Arg Gly Glu Leu Glu Ser Ile Leu Trp Gln Ala
 305 310 315 320
 Glu Gln Pro Val Ser Gln Thr Thr Thr Leu Pro Gln Lys Val Leu Asp
 325 330 335
 Ser Thr Pro Val Leu Asp Ser Val Leu Ser Glu Ser Leu Arg Leu Thr
 340 345 350
 Ala Ala Pro Phe Ile Thr Arg Glu Val Val Val Asp Leu Ala Met Pro
 355 360 365
 Met Ala Asp Gly Arg Glu Phe Asn Leu Arg Arg Gly Asp Arg Leu Leu
 370 375 380
 Leu Phe Pro Phe Leu Ser Pro Gln Arg Asp Pro Glu Ile Tyr Thr Asp
 385 390 395 400
 Pro Glu Val Phe Lys Tyr Asn Arg Phe Leu Asn Pro Asp Gly Ser Glu
 405 410 415
 Lys Lys Asp Phe Tyr Lys Asp Gly Lys Arg Leu Lys Asn Tyr Asn Met
 420 425 430
 Pro Trp Gly Ala Gly His Asn His Cys Leu Gly Arg Ser Tyr Ala Val
 435 440 445
 Asn Ser Ile Lys Gln Phe Val Phe Leu Val Leu Val His Leu Asp Leu
 450 455 460
 Glu Leu Ile Asn Ala Asp Val Glu Ile Pro Glu Phe Asp Leu Ser Arg
 465 470 475 480
 Tyr Gly Phe Gly Leu Met Gln Pro Glu His Asp Val Pro Val Arg Tyr
 485 490 495
 Arg Ile Arg Pro
 500

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "PRIMER/SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GACAAGGACC ACATGTGCAG TGTC

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "PRIMER/SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CTGTGTGGCC CACAGCTGCA GCAC